

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT: Hein, Mich B.
Hiatt, Andrew C.
Ma, Julian K.C.

10

(ii) TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ASSEMBLED
SECRETORY ANTIBODIES

(iii) NUMBER OF SEQUENCES: 26

15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
(B) STREET: 10666 North Torrey Pines Road, TPC-8
(C) CITY: La Jolla
(D) STATE: California
(E) COUNTRY: US
(F) ZIP: 92037

20

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE: 03-MAY-1996
(C) CLASSIFICATION:

30

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/591,823
(B) FILING DATE: 02-OCT-1990

35

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/427,765
(B) FILING DATE: 27-OCT-1989

40

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Logan, April C.
(B) REGISTRATION NUMBER: 33,950
(C) REFERENCE/DOCKET NUMBER: 184.2

45

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (619) 554-2937
(B) TELEFAX: (619) 554-6312

50

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

15

CCTTGACCGT AAGACATG

18

(2) INFORMATION FOR SEQ ID NO:2:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

30

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

35

AATTCATGTC TTACGGTCAA GG

22

(2) INFORMATION FOR SEQ ID NO:3:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGTGAAAACC ATATTGAATT CCACCAATAC AAA

33

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATTTAGCACA ACATCCATGT CGACGAATTC AATCCAAAAA AGCAT

45

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGGAGCTGG TGGTGAATT CGTCGACCTT TGTCTCTAAC AC

42

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5

CCATCCCATG GTTGAATTCA GTGTCGTCAG

30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

25

CTGCAACTGG ACCTGCATGT CGACGAATTC AGCTCCTGAC AGGAG

45

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

45

CCTGTAGGAC CAGAGGAATT CGTCGACACT GGGATTATTT AC

42

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 75 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

10 GAATTCATTC AAGAATAGTT CAAACAAGAA GATTACAAAC TATCAATTC ATACACAATA 60
TAAACGATTA AAAGA 75

15 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: amino acid
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

30 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30

35 Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45

40 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
65 70 75 80

45 Ser Leu Asp Leu Lys Arg Asp Val Val Leu
85 90

(2) INFORMATION FOR SEQ ID NO:11:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

10 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30
15 Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60
20 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
65 70 75 80
Ser Leu Asp Leu Lys Arg Glu Val Glu Leu
25 85 90

(2) INFORMATION FOR SEQ ID NO:12:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

40 Met Glu Leu Asp Leu Ser Leu Pro Leu Ser Gly Ala Ala Gly Gly Thr
1 5 10 15

45 (2) INFORMATION FOR SEQ ID NO:13:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

31

ACCAGATCTA TGGAAATGGAC CTGGGTTTTT C

10

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: DNA (genomic)

20

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

30

CCCAAGCTTG GTTTTGGAGA TGGTTTTCTC

30

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

40

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

31

GATAAGCTTG GTCCTACTCC TCCTCCTCCT A

50

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

30

AATCTCGAGT CAGTAGCAGA TGCCATCTCC

15 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

30

GGAAAGCTTT GTACATATGC AAGGCTTACA

35 (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

45 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

23

GATCTATGGC TCTCTTCTTG CTC

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AATTCTTATT CCGCACTCTG CACTGC

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3517 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

40	GGCCGGGGTT ACGGGCTGGC CAGCAGGCTG TGCCCCGAG TCCGGTCAGC AGGAGGGGAA	60
	GAAGTGGCCT AAAATCTCTC CCGCATCGGC AGCCCAGGCC TAGTGCCCTA CCAGCCACCA	120
45	GCCATGGCTC TCTTCTTGCT CACCTGCCTG CTGGCTGTCT TTTCAGCGGC CACGGCACAA	180
	AGCTCCTTAT TGGGTCCCAG CTCCATATTT GGTCCCGGGG AGGTGAATGT TTTGGAAGGC	240
	GACTCGGTGT CCATCACATG CTACTACCCA ACAACCTCCG TCACCCGGCA CAGCCGGAAG	300
50	TTCTGGTGCC GGAAGAGGA GAGCGGCCGC TGCCTGACGC TTGCCTCGAC CGGCTACACG	360
	TCCCAGGAAT ACTCCGGGAG AGGCAAGCTC ACCGACTTCC CTGATAAAGG GGAGTTTGTG	420

	GTGACTGTTG ACCAACTCAC CCAGAACGAC TCAGGGAGCT ACAAGTGTGG CGTGGGAGTC	480
	AACGGCCGTG GCCTGGACTT CGGTGTCAAC GTGCTGGTCA GCCAGAAGCC AGAGCCTGAT	540
5	GACGTTGTTT ACAAACAATA TGAGAGTTAT ACAGTAACCA TCACCTGCCC TTTCACATAT	600
	GCGACTAGGC AACTAAAGAA GTCCTTTTAC AAGGTGGAAG ACGGGGAAC TGTACTCATC	660
	ATTGATTCCA GCAGTAAGGA GGCAAAGGAC CCCAGGTATA AGGGCAGAAT AACGTTGCAG	720
10	ATCCAAAGTA CCACAGCAAA AGAATTCACA GTCACCATCA AGCATTTGCA GCTCAATGAT	780
	GCTGGGCAGT ATGTCTGCCA GAGTGGAAGC GACCCCACTG CTGAAGAACA GAACGTTGAC	840
15	CTCCGACTGC TAACTCCTGG TCTGCTCTAT GGAAACCTGG GGGGCTCGGT GACCTTTGAA	900
	TGTGCCCTGG ACTCTGAAGA CGCAAACGCG GTAGCATCCT TGCGCCAGGT TAGGGGTGGC	960
	AATGTGGTCA TTGACAGCCA GGGGACAATA GATCCAGCCT TCGAGGGCAG GATCCTGTTC	1020
20	ACCAAGGCTG AGAACGGCCA CTTCAAGTGA GTGATCGCAG GCCTGAGGAA GGAAGACACA	1080
	GGGAACTATC TGTGCGGAGT CCAGTCCAAT GGTCACTCTG GGGATGGGCC CACCCAGCTT	1140
25	CGGCAACTCT TCGTCAATGA AGAGATCGAC GTGTCCCGCA GCCCCCCTGT GTTGAAGCGC	1200
	TTTCCAGGAG GCTCCGTGAC CATAAGCTGC CCCTACAACC CGAAGAGAAG CGACAGCCAC	1260
	CTGCAGCTGT ATCTCTGGGA AGGGAGTCAA ACCCGCCATC TGCTGGTGA CAGCGGCGAG	1320
30	GGGCTGGTTC AGAAAGACTA CACAGGCAGG CTGGCCCTGT TCGAAGAGCC TGGCAATGGC	1380
	ACCTTCTCAG TCGTCCTCAA CCAGCTCACT GCCGAGGATG AAGGCTTCTA CTGGTGTGTC	1440
35	AGCGATGACG ATGAGTCCCT GACGACTTCG GTGAAGCTCC AGATCGTTGA CGGAGAACCA	1500
	AGCCCCACGA TCGACAAGTT CACTGCTGTG CAGGGAGAGC CTGTTGAGAT CACCTGCCAC	1560
40	TTCCCATGCA AATACTTCTC CTCCGAGAAG TACTGGTGCA AGTGAATGA CCATGGCTGC	1620
	GAGGACCTGC CCACTAAGCT CAGCTCCAGC GGCGACCTTG TGAAATGCAA CAACAACCTG	1680
	GTCCTCACCC TGACCTTGGA CTCGGTCAGC GAAGATGACG AGGGCTGGTA CTGGTGTGGC	1740
45	GCGAAAGACG GGCACGAGTT TGAAGAGGTT GCGGCCGTCA GGGTGGAGCT GACAGAGCCA	1800
	GCCAAGGTAG CTGTGAGGCC AGCCAGGTA CCTGTGACCC CAGCCAAGGC AGCCCCGCG	1860
	CCTGCTGAGG AGAAGGCCAA GCGCGGTGTC CCAGTGCCCA GGAGAAGGCA GTGGTACCCA	1920
50	TTGTCAAGGA AGCTGAGAAC AAGTTGTCCA GAACCTCGGC TCCTTGCGGA GGAGGTAGCA	1980
	GTGCAGAGTG CGGAAGACCC AGCCAGTGGG AGCAGAGCGT CTGTGGATGC CAGCAGTGCT	2040

	TCGGGACAAA GCGGGAGTGC CAAAGTACTG ATCTCCACCC TGGTGCCCTT GGGGCTGGTG	2100
	CTGGCAGCGG GGGCCATGGC CGTGGCCATA GCCAGAGCCC GGCACAGGAG GAACGTGGAC	2160
5	CGAGTTTCCA TCGGAAGCTA CAGGACAGAC ATTAGCATGT CAGACTTGGA GAACTCCAGG	2220
	GAGTTCGGAG CCATTGACAA CCCAAGCGCC TGCCCCGATG CCCGGGAGAC GGCCCTCGGA	2280
	GGAAAGGATG AGTTAGCGAC GGCCACCGAG AGCACC GTGG AGATTGAGGA GCCCAAGAAG	2340
10	GCAAAACGGT CATCCAAGGA AGAAGCCGAC CTGGCCTACT CAGCTTTCCT GCTCCAATCC	2400
	AACACCATAG CTGCTGAGCA CCAAGATGGC CCCAAGGAGG CCTAGGCACA GCCGGCCACC	2460
15	GCCGCCGCCG CCACCSCCGC CGCCGCCGCC ACCTGTGAAA ATCACCTTCC AGAATCACGT	2520
	TGATCCTCGG GGTCCCCAGA GCCGGGGGCT CAACCGCCCT GCACCCCCCA TGTCCCCACC	2580
	ACCTAAACTT CCCTACCTGT GCCCAGAGGT GTGCTGGTCC CCTCCTCCAC GGCATCCAGG	2640
20	CCTGGCTCAA TGTTCCCGTT GGGGTGGGGG TGTGAGGGGT TCCTACTTGC AGCCCGGTTC	2700
	TCCCCAGAGA AGCTAAGGAT CCAGGTCCTG AGGGAGGGGC CTCTCGAAGG CAGACAGACC	2760
25	AGAGAGGGGG GAGGAGCCCT TGGATGGGAG GCCAGAGGCG CTTTCCGGCC ACCCCCTCCC	2820
	TCCCTGCCCC CACCCTCCTT CCTTCATTCA AAAGTCCCAG TGGCTGCTGC CTAGGGTCCA	2880
	GGCGCTGGCC GCACGCCTCC TCGAAGCCGT TGTGCAAACA TCACTGGAGG AAGCCAGGGC	2940
30	TCCTCCCGGG CTGTGTATCC TCACTCAGGC ATCCTGTCTT CCCCAGTATC AGGAGATGTC	3000
	AAGCCTCTGA AGGCTGTGTG CCCTGGGCGT GTCTGCAAGT CACCCCAGAC ACATGTTCTC	3060
35	GCCATTTTAC AGATGAGAAC ACTGAGGTG TACTCAAGGG CACCCTGCGA GATGGAGCAA	3120
	CAGCAAATA GATGGGCTTC TGCTGTCTCT TTGGCCAGAG GTCTCTCCAC AGGAGCCCCCT	3180
	GCCCCGTAG GAAGCAGAGT TTTAGAACAT GGAAGAAGAA GAGGGGGATG GCCCTGGACG	3240
40	CTGACCTCTC CCAAGCCCCC ACGGGGGAAA AGGCCCCCTC CTTTTCTGTC ACTCTCGGGG	3300
	ACCTGCGGAG TTGAGCATTG GTGCCCCGTG TGTCTGAAGA GTTCCCAGTG GAAAGAAGAA	3360
45	AAGAGGGTGT TTGTCAAGTG CGGGGAGGGC CTGATCCCCA GACAGCTGAA GTTTAAGGTC	3420
	CTTGTCCTTG TGAGCTTTAA CCAGCACCTC CGGGCTGACC CTTGCTAACA CATCAGAAAT	3480
50	GTGATTTAAT CATTAAACAT TGTGATTGCC ACTGGGA	3517

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1875 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

10

(iv) ANTI-SENSE: NO

(ix) FEATURE:

15

(A) NAME/KEY: CDS

(B) LOCATION: 1..1875

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

20

ATG GCT CTC TTC TTG CTC ACC TGC CTG CTG GCT GTC TTT TCA GCG GCC 48
 Met Ala Leu Phe Leu Leu Thr Cys Leu Leu Ala Val Phe Ser Ala Ala
 1 5 10 15

25

ACG GCA CAA AGC TCC TTA TTG GGT CCC AGC TCC ATA TTT GGT CCC GGG 96
 Thr Ala Gln Ser Ser Leu Leu Gly Pro Ser Ser Ile Phe Gly Pro Gly
 20 25 30

30

GAG GTG AAT GTT TTG GAA GGC GAC TCG GTG TCC ATC ACA TGC TAC TAC 144
 Glu Val Asn Val Leu Glu Gly Asp Ser Val Ser Ile Thr Cys Tyr Tyr
 35 40 45

35

CCA ACA ACC TCC GTC ACC CGG CAC AGC CGG AAG TTC TGG TGC CGG GAA 192
 Pro Thr Thr Ser Val Thr Arg His Ser Arg Lys Phe Trp Cys Arg Glu
 50 55 60

40

GAG GAG AGC GGC CGC TGC GTG ACG CTT GCC TCG ACC GGC TAC ACG TCC 240
 Glu Glu Ser Gly Arg Cys Val Thr Leu Ala Ser Thr Gly Tyr Thr Ser
 65 70 75 80

45

CAG GAA TAC TCC GGG AGA GGC AAG CTC ACC GAC TTC CCT GAT AAA GGG 288
 Gln Glu Tyr Ser Gly Arg Gly Lys Leu Thr Asp Phe Pro Asp Lys Gly
 85 90 95

50

GAG TTT GTG GTG ACT GTT GAC CAA CTC ACC CAG AAC GAC TCA GGG AGC 336
 Glu Phe Val Val Thr Val Asp Gln Leu Thr Gln Asn Asp Ser Gly Ser
 100 105 110

TAC AAG TGT GGC GTG GGA GTC AAC GGC CGT GGC CTG GAC TTC GGT GTC 384
 Tyr Lys Cys Gly Val Gly Val Asn Gly Arg Gly Leu Asp Phe Gly Val
 115 120 125

AAC GTG CTG GTC AGC CAG AAG CCA GAG CCT GAT GAC GTT GTT TAC AAA 432
 Asn Val Leu Val Ser Gln Lys Pro Glu Pro Asp Asp Val Val Tyr Lys

	130	135	140	
5	CAA TAT GAG AGT TAT ACA GTA ACC ATC ACC TGC CCT TTC ACA TAT GCG Gln Tyr Glu Ser Tyr Thr Val Thr Ile Thr Cys Pro Phe Thr Tyr Ala 145 150 155 160	480		
10	ACT AGG CAA CTA AAG AAG TCC TTT TAC AAG GTG GAA GAC GGG GAA CTT Thr Arg Gln Leu Lys Lys Ser Phe Tyr Lys Val Glu Asp Gly Glu Leu 165 170 175	528		
15	GTA CTC ATC ATT GAT TCC AGC AGT AAG GAG GCA AAG GAC CCC AGG TAT Val Leu Ile Ile Asp Ser Ser Ser Lys Glu Ala Lys Asp Pro Arg Tyr 180 185 190	576		
20	AAG GGC AGA ATA ACG TTG CAG ATC CAA AGT ACC ACA GCA AAA GAA TTC Lys Gly Arg Ile Thr Leu Gln Ile Gln Ser Thr Thr Ala Lys Glu Phe 195 200 205	624		
25	ACA GTC ACC ATC AAG CAT TTG CAG CTC AAT GAT GCT GGG CAG TAT GTC Thr Val Thr Ile Lys His Leu Gln Leu Asn Asp Ala Gly Gln Tyr Val 210 215 220	672		
30	TGC CAG AGT GGA AGC GAC CCC ACT GCT GAA GAA CAG AAC GTT GAC CTC Cys Gln Ser Gly Ser Asp Pro Thr Ala Glu Glu Gln Asn Val Asp Leu 225 230 235 240	720		
35	CGA CTG CTA ACT CCT GGT CTG CTC TAT GGA AAC CTG GGG GGC TCG GTG Arg Leu Leu Thr Pro Gly Leu Leu Tyr Gly Asn Leu Gly Gly Ser Val 245 250 255	768		
40	ACC TTT GAA TGT GCC CTG GAC TCT GAA GAC GCA AAC GCG GTA GCA TCC Thr Phe Glu Cys Ala Leu Asp Ser Glu Asp Ala Asn Ala Val Ala Ser 260 265 270	816		
45	TTG CGC CAG GTT AGG GGT GGC AAT GTG GTC ATT GAC AGC CAG GGG ACA Leu Arg Gln Val Arg Gly Gly Asn Val Val Ile Asp Ser Gln Gly Thr 275 280 285	864		
50	ATA GAT CCA GCC TTC GAG GGC AGG ATC CTG TTC ACC AAG GCT GAG AAC Ile Asp Pro Ala Phe Glu Gly Arg Ile Leu Phe Thr Lys Ala Glu Asn 290 295 300	912		
	GGC CAC TTC AGT GTA GTG ATC GCA GGC CTG AGG AAG GAA GAC ACA GGG Gly His Phe Ser Val Val Ile Ala Gly Leu Arg Lys Glu Asp Thr Gly 305 310 315 320	960		
	AAC TAT CTG TGC GGA GTC CAG TCC AAT GGT CAG TCT GGG GAT GGG CCC Asn Tyr Leu Cys Gly Val Gln Ser Asn Gly Gln Ser Gly Asp Gly Pro 325 330 335	1008		
	ACC CAG CTT CGG CAA CTC TTC GTC AAT GAA GAG ATC GAC GTG TCC CGC Thr Gln Leu Arg Gln Leu Phe Val Asn Glu Glu Ile Asp Val Ser Arg 340 345 350	1056		

	AGC CCC CCT GTG TTG AAG GGC TTT CCA GGA GGC TCC GTG ACC ATA CGC	1104
	Ser Pro Pro Val Leu Lys Gly Phe Pro Gly Gly Ser Val Thr Ile Arg	
	355 360 365	
5	TGC CCC TAC AAC CCG AAG AGA AGC GAC AGC CAC CTG CAG CTG TAT CTC	1152
	Cys Pro Tyr Asn Pro Lys Arg Ser Asp Ser His Leu Gln Leu Tyr Leu	
	370 375 380	
10	TGG GAA GGG AGT CAA ACC CGC CAT CTG CTG GTG GAC AGC GGC GAG GGG	1200
	Trp Glu Gly Ser Gln Thr Arg His Leu Leu Val Asp Ser Gly Glu Gly	
	385 390 395 400	
15	CTG GTT CAG AAA GAC TAC ACA GGC AGG CTG GCC CTG TTC GAA GAG CCT	1248
	Leu Val Gln Lys Asp Tyr Thr Gly Arg Leu Ala Leu Phe Glu Glu Pro	
	405 410 415	
20	GGC AAT GGC ACC TTC TCA GTC GTC CTC AAC CAG CTC ACT GCC GAG GAT	1296
	Gly Asn Gly Thr Phe Ser Val Val Leu Asn Gln Leu Thr Ala Glu Asp	
	420 425 430	
25	GAA GGC TTC TAC TGG TGT GTC AGC GAT GAC GAT GAG TCC CTG ACG ACT	1344
	Glu Gly Phe Tyr Trp Cys Val Ser Asp Asp Asp Glu Ser Leu Thr Thr	
	435 440 445	
30	TCG GTG AAG CTC CAG ATC GTT GAC GGA GAA CCA AGC CCC ACG ATC GAC	1392
	Ser Val Lys Leu Gln Ile Val Asp Gly Glu Pro Ser Pro Thr Ile Asp	
	450 455 460	
35	AAG TTC ACT GCT GTG CAG GGA GAG CCT GTT GAG ATC ACC TGC CAC TTC	1440
	Lys Phe Thr Ala Val Gln Gly Glu Pro Val Glu Ile Thr Cys His Phe	
	465 470 475 480	
40	CCA TGC AAA TAC TTC TCC TCC GAG AAG TAC TGG TGC AAG TGG AAT GAC	1488
	Pro Cys Lys Tyr Phe Ser Ser Glu Lys Tyr Trp Cys Lys Trp Asn Asp	
	485 490 495	
45	CAT GGC TGC GAG GAC CTG CCC ACT AAG CTC AGC TCC AGC GGC GAC CTT	1536
	His Gly Cys Glu Asp Leu Pro Thr Lys Leu Ser Ser Ser Gly Asp Leu	
	500 505 510	
50	GTG AAA TGC AAC AAC AAC CTG GTC CTC ACC CTG ACC TTG GAC TCG GTC	1584
	Val Lys Cys Asn Asn Asn Leu Val Leu Thr Leu Thr Leu Asp Ser Val	
	515 520 525	
55	AGC GAA GAT GAC GAG GGC TGG TAC TGG TGT GGC GCG AAA GAC GGG CAC	1632
	Ser Glu Asp Asp Glu Gly Trp Tyr Trp Cys Gly Ala Lys Asp Gly His	
	530 535 540	
60	GAG TTT GAA GAG GTT GCG GCC GTC AGG GTG GAG CTG ACA GAG CCA GCC	1680
	Glu Phe Glu Glu Val Ala Ala Val Arg Val Glu Leu Thr Glu Pro Ala	
	545 550 555 560	
65	AAG GTA GCT GTC GAG CCA GCC AAG GTA CCT GTC GAC CCA GCC AAG GCA	1728

Lys Val Ala Val Glu Pro Ala Lys Val Pro Val Asp Pro Ala Lys Ala
 565 570 575

5 GCC CCC GCG CCT GCT GAG GAG AAG GCC AAG GCG CGG TGC CCA GTG CCC 1776
 Ala Pro Ala Pro Ala Glu Glu Lys Ala Lys Ala Arg Cys Pro Val Pro
 580 585 590

10 AGG AGA AGG CAG TGG TAC CCA TTG TCA AGG AAG CTG AGA ACA AGT TGT 1824
 Arg Arg Arg Gln Trp Tyr Pro Leu Ser Arg Lys Leu Arg Thr Ser Cys
 595 600 605

15 CCA GAA CCT CGG CTC CTT GCG GAG GAG GTA GCA GTG CAG AGT GCG GAA 1872
 Pro Glu Pro Arg Leu Leu Ala Glu Glu Val Ala Val Gln Ser Ala Glu
 610 615 620

TA
 625

20 (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 624 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Ala Leu Phe Leu Leu Thr Cys Leu Leu Ala Val Phe Ser Ala Ala
 1 5 10 15

35 Thr Ala Gln Ser Ser Leu Leu Gly Pro Ser Ser Ile Phe Gly Pro Gly
 20 25 30

Glu Val Asn Val Leu Glu Gly Asp Ser Val Ser Ile Thr Cys Tyr Tyr
 35 40 45

40 Pro Thr Thr Ser Val Thr Arg His Ser Arg Lys Phe Trp Cys Arg Glu
 50 55 60

45 Glu Glu Ser Gly Arg Cys Val Thr Leu Ala Ser Thr Gly Tyr Thr Ser
 65 70 75 80

Gln Glu Tyr Ser Gly Arg Gly Lys Leu Thr Asp Phe Pro Asp Lys Gly
 85 90 95

50 Glu Phe Val Val Thr Val Asp Gln Leu Thr Gln Asn Asp Ser Gly Ser
 100 105 110

Tyr Lys Cys Gly Val Gly Val Asn Gly Arg Gly Leu Asp Phe Gly Val

	115	120	125
	Asn Val Leu Val Ser Gln Lys Pro Glu Pro Asp Asp Val Val Tyr Lys		
	130	135	140
5	Gln Tyr Glu Ser Tyr Thr Val Thr Ile Thr Cys Pro Phe Thr Tyr Ala		
	145	150	155 160
10	Thr Arg Gln Leu Lys Lys Ser Phe Tyr Lys Val Glu Asp Gly Glu Leu		
		165	170 175
	Val Leu Ile Ile Asp Ser Ser Ser Lys Glu Ala Lys Asp Pro Arg Tyr		
		180	185 190
15	Lys Gly Arg Ile Thr Leu Gln Ile Gln Ser Thr Thr Ala Lys Glu Phe		
		195	200 205
	Thr Val Thr Ile Lys His Leu Gln Leu Asn Asp Ala Gly Gln Tyr Val		
		210	215 220
20	Cys Gln Ser Gly Ser Asp Pro Thr Ala Glu Glu Gln Asn Val Asp Leu		
		225	230 235 240
25	Arg Leu Leu Thr Pro Gly Leu Leu Tyr Gly Asn Leu Gly Gly Ser Val		
		245	250 255
	Thr Phe Glu Cys Ala Leu Asp Ser Glu Asp Ala Asn Ala Val Ala Ser		
		260	265 270
30	Leu Arg Gln Val Arg Gly Gly Asn Val Val Ile Asp Ser Gln Gly Thr		
		275	280 285
	Ile Asp Pro Ala Phe Glu Gly Arg Ile Leu Phe Thr Lys Ala Glu Asn		
		290	295 300
35	Gly His Phe Ser Val Val Ile Ala Gly Leu Arg Lys Glu Asp Thr Gly		
		305	310 315 320
40	Asn Tyr Leu Cys Gly Val Gln Ser Asn Gly Gln Ser Gly Asp Gly Pro		
		325	330 335
	Thr Gln Leu Arg Gln Leu Phe Val Asn Glu Glu Ile Asp Val Ser Arg		
		340	345 350
45	Ser Pro Pro Val Leu Lys Gly Phe Pro Gly Gly Ser Val Thr Ile Arg		
		355	360 365
	Cys Pro Tyr Asn Pro Lys Arg Ser Asp Ser His Leu Gln Leu Tyr Leu		
		370	375 380
50	Trp Glu Gly Ser Gln Thr Arg His Leu Leu Val Asp Ser Gly Glu Gly		
		385	390 395 400

	Leu Val Gln Lys Asp Tyr Thr Gly Arg Leu Ala Leu Phe Glu Glu Pro	405	410	415
5	Gly Asn Gly Thr Phe Ser Val Val Leu Asn Gln Leu Thr Ala Glu Asp	420	425	430
	Glu Gly Phe Tyr Trp Cys Val Ser Asp Asp Asp Glu Ser Leu Thr Thr	435	440	445
10	Ser Val Lys Leu Gln Ile Val Asp Gly Glu Pro Ser Pro Thr Ile Asp	450	455	460
	Lys Phe Thr Ala Val Gln Gly Glu Pro Val Glu Ile Thr Cys His Phe	465	470	475
15	Pro Cys Lys Tyr Phe Ser Ser Glu Lys Tyr Trp Cys Lys Trp Asn Asp	485	490	495
	His Gly Cys Glu Asp Leu Pro Thr Lys Leu Ser Ser Ser Gly Asp Leu	500	505	510
20	Val Lys Cys Asn Asn Asn Leu Val Leu Thr Leu Thr Leu Asp Ser Val	515	520	525
	Ser Glu Asp Asp Glu Gly Trp Tyr Trp Cys Gly Ala Lys Asp Gly His	530	535	540
25	Glu Phe Glu Glu Val Ala Ala Val Arg Val Glu Leu Thr Glu Pro Ala	545	550	555
30	Lys Val Ala Val Glu Pro Ala Lys Val Pro Val Asp Pro Ala Lys Ala	565	570	575
	Ala Pro Ala Pro Ala Glu Glu Lys Ala Lys Ala Arg Cys Pro Val Pro	580	585	590
35	Arg Arg Arg Gln Trp Tyr Pro Leu Ser Arg Lys Leu Arg Thr Ser Cys	595	600	605
40	Pro Glu Pro Arg Leu Leu Ala Glu Glu Val Ala Val Gln Ser Ala Glu	610	615	620

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- 50
- (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

10

GATCTATGAA GACCCACCTG CTT

23

(2) INFORMATION FOR SEQ ID NO:24:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

25

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

30

AATTCTTAGA CAGGGTAGCA AGA

23

(2) INFORMATION FOR SEQ ID NO:25:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 480 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

45

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

50

(B) LOCATION: 1..480

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

	ATG AAG ACC CAC CTG CTT CTC TGG GGA GTC CTC GCC ATT TTT GTT AAG	48
	Met Lys Thr His Leu Leu Leu Trp Gly Val Leu Ala Ile Phe Val Lys	
	1 5 10 15	
5	GTT GTC CTT GTA ACA GGT GAC GAC GAA GCG ACC ATT CTT GCT GAC AAC	96
	Val Val Leu Val Thr Gly Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn	
	20 25 30	
10	AAA TGC ATG TGT ACC CGA GTT ACC TCT AAA ATC ATC CCT TCC ACC GAG	144
	Lys Cys Met Cys Thr Arg Val Thr Ser Lys Ile Ile Pro Ser Thr Glu	
	35 40 45	
15	GAT CCT AAT GAG GAC ATT GTG GAG AGA AAT ATC CGA ATT GTT GTC CCT	192
	Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Val Val Pro	
	50 55 60	
20	TTG AAC AAC AGG GAG AAT ATC TCT GAT CCC ACC TCC CCA CTG AGA AGG	240
	Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg	
	65 70 75 80	
25	AAC TTT GTA TAC CAT TTG TCA GAC GTC TGT AAG AAA TGC GAT CCT GTG	288
	Asn Phe Val Tyr His Leu Ser Asp Val Cys Lys Lys Cys Asp Pro Val	
	85 90 95	
30	GAA GTG GAG CTG GAA GAT CAG GTT GTT ACT GCC ACC CAG AGC AAC ATC	336
	Glu Val Glu Leu Glu Asp Gln Val Thr Ala Thr Gln Ser Asn Ile	
	100 105 110	
35	TGC AAT GAA GAC GAT GGT GTT CCT GAG ACC TGC TAC ATG TAT GAC AGA	384
	Cys Asn Glu Asp Asp Gly Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg	
	115 120 125	
40	AAC AAG TGC TAT ACC ACT ATG GTC CCA CTT AGG TAT CAT GGT GAG ACC	432
	Asn Lys Cys Tyr Thr Thr Met Val Pro Leu Arg Tyr His Gly Glu Thr	
	130 135 140	
45	AAA ATG GTG CAA GCA GCC TTG ACC CCC GAT TCT TGC TAC CCT GAC TA	480
	Lys Met Val Gln Ala Ala Leu Thr Pro Asp Ser Cys Tyr Pro Asp	
	145 150 155 160	

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 159 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Lys Thr His Leu Leu Leu Trp Gly Val Leu Ala Ile Phe Val Lys

	1	5	10	15
	Val Val Leu Val Thr Gly Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn	20	25	30
5	Lys Cys Met Cys Thr Arg Val Thr Ser Lys Ile Ile Pro Ser Thr Glu	35	40	45
10	Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Val Val Pro	50	55	60
	Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg	65	70	75
15	Asn Phe Val Tyr His Leu Ser Asp Val Cys Lys Lys Cys Asp Pro Val	85	90	95
	Glu Val Glu Leu Glu Asp Gln Val Val Thr Ala Thr Gln Ser Asn Ile	100	105	110
20	Cys Asn Glu Asp Asp Gly Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg	115	120	125
25	Asn Lys Cys Tyr Thr Thr Met Val Pro Leu Arg Tyr His Gly Glu Thr	130	135	140
	Lys Met Val Gln Ala Ala Leu Thr Pro Asp Ser Cys Tyr Pro Asp	145	150	155